



Multi-model inference in comparative phylogeography: an integrative approach based on multiple lines of evidence

Rosane G. Collevatti^{1*}, Levi C. Terribile², José A. F. Diniz-Filho³ and Matheus S. Lima-Ribeiro²

¹ Laboratório de Genética & Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, Brazil

² Laboratório de Macroecologia, Universidade Federal de Goiás, Jataí, Brazil

³ Departamento de Ecologia, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, Brazil

Edited by:

Toby Pennington, Royal Botanic Garden Edinburgh, UK

Reviewed by:

David Vieites, Spanish Research Council – Consejo Superior de Investigaciones Científicas, Spain
Ana C. Carnaval, City College of New York, USA

*Correspondence:

Rosane G. Collevatti, Laboratório de Genética & Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal de Goiás, C. P. 131, Campus II, Setor Samambaia, 74001-970 Goiânia, GO, Brazil
e-mail: rosanegc68@hotmail.com

Comparative phylogeography has its roots in classical biogeography and, historically, relies on a pattern-based approach. Here, we present a model-based framework for comparative phylogeography. Our framework was initially developed for statistical phylogeography based on a multi-model inference approach, by coupling ecological niche modeling, coalescent simulation and direct spatio-temporal reconstruction of lineage diffusion using a relaxed random walk model. This multi-model inference framework is particularly useful to investigate the complex dynamics and current patterns in genetic diversity in response to processes operating on multiple taxonomic levels in comparative phylogeography. In addition, because of the lack, or incompleteness of fossil record, the understanding of the role of biogeographical events (vicariance and dispersal routes) in most regions worldwide is barely known. Thus, we believe that the expansion of that framework for multiple species under a comparative approach may give clues on genetic legacies in response to Quaternary climate changes and other biogeographical processes.

Keywords: biogeography, coalescence, palaeodistribution modeling, Quaternary climatic changes, vicariance

INTRODUCTION

Comparative phylogeography has historically been derived from classical historical biogeography, whereby common patterns in lineage distribution within multiple taxa are often explained by vicariant events shared by all taxa (Bermingham and Moritz, 1998; Avise, 2000). However, because similar population genetic structures may arise under different demographic processes, the conventional method based on narrative descriptions and pattern interpretation derived from historical biogeography often results in dubious or indistinguishable historical demographic or vicariant processes. Thus, recovering the true demographic history of species is critical for understanding microevolutionary processes and the spatial context of lineage divergences (Knowles and Maddison, 2002), but this approach should still be expanded to the context of comparative phylogeography.

In this context, one of the major challenges in the emergent field of statistical phylogeography is to set up demographic scenarios independently of gene trees, which should help to define alternative hypotheses for temporal and geographical aspects of species dynamics and, consequently, establish relevant phylogeographical inferences (Knowles et al., 2007; Knowles, 2009; Carstens and Knowles, 2010). Recently, phylogeographers began to explore multiple lines of evidence obtained from advances in geology, paleontology, palaeopalynology, and ecology to guide choices about species dynamics across time and space, such as: (1) fossil and archeological records, including ancient DNA, as direct empirical evidence (Lagerholm et al., 2014); (2) palaeoclimatological and palaeovegetational reconstructions or general patterns of species distribution based on floristic records, as

indirect evidence (e.g., Collevatti et al., 2012a,b); (3) palaeodistribution modeling, the historical extension of a model-based approach increasingly applied to macroecological and palaeobiological questions by coupling the theory of ecological niche with palaeoclimatic simulations (e.g., Carnaval and Moritz, 2008); and (4) a combination of two or more of these approaches (e.g., Metcalf et al., 2014). Due to incompleteness of the fossil record (Paul, 2009), as well as the coarse nature of the temporal and spatial resolution of palaeoecological reconstructions, the use of ecological niche models has been an accessible and efficient tool to incorporate explicit spatio-temporal information into analyses of gene trees. Such links between statistical phylogeography and macroecology are indeed an emergent approach of current phylogeographical analyses (Richards et al., 2007; Collevatti et al., 2013).

More recently, models based on Approximate Bayesian Computation (ABC) implemented in MTML-msBayes (Huang et al., 2011) have been used to test simultaneous divergence in comparative phylogeography, taking into account the stochastic variance in coalescence processes underlying multiple co-distributed lineages, thus providing general biogeographic explanations for phylogeographical patterns (e.g., Chan et al., 2011; Bell et al., 2012). Here we discuss and propose perspectives for expansion of this new framework of comparative analyses by exploring a multi-model inference approach, centered on the recent advances of statistical phylogeography (see Collevatti et al., 2012a, 2013). We first present the multi-model approach and its components and then discuss how it can be expanded to infer processes in a comparative fashion.

COUPLING STATISTICAL PHYLOGEOGRAPHY AND MACROECOLOGICAL APPROACHES IN A MULTI-MODEL INFERENCE FRAMEWORK

The reasoning for inferring processes from a multitude of models as an alternative to null-hypothesis significance testing comes from the common need to search general explanations for observed patterns from unknown explicit causal mechanisms (Stephens et al., 2006). Multi-model inference assumes that the most effective processes causing a given observed pattern should be inferred from the best-fitted model describing empirical data (Grueber et al., 2011). However, models used in practice and the inferred processes are limited to the available information. Paul Velleman clearly synthesizes this limiting factor:

“A model for data, no matter how elegant, or correctly derived, must be discarded or revised if it does not fit the data or when new or better data are found and it fails to fit them” (Velleman, 2008, p. 4).

Thus, uncertainties from multiple sources of evidence should be exhaustively explored to reach the most possible realistic and less changeable inferences, and eventually the level of uncertainty can be so high that actually many alternative models become available, allowing a multi model-based approach (Millington and Perry, 2011).

In the multi-model approach for phylogeography, uncertainties can be investigated (i) within the set of alternative demographic scenarios being considered (for instance, exploring the impact of alternative methods and climatic simulations in ecological niche modeling), (ii) during the coalescent simulation steps (when estimating genetic and dispersal parameters), and lastly (iii) at the model selection stage (given multiple available selection criteria, e.g., Akaike's Information Criteria and likelihood estimates; Csilléry et al., 2010). Within the sub discipline of comparative phylogeography, we further propose to explore uncertainties from multiple phylogeographical hypotheses of the studied species.

SPECIES NICHE, PALAEO DISTRIBUTION MODELING, AND THE ALTERNATIVE DEMOGRAPHICAL HYPOTHESIS

Ecological niche modeling (ENM) has allowed exploration of the geographic context of species dynamics through the past by hind casting suitable climatic conditions (an n -dimensional space of climatic variables) using palaeoclimatic scenarios (Martínez-Meyer et al., 2004; Nogués-Bravo, 2009). ENMs deal with the geographic context independent of gene trees, and predict species' potential distribution over different time periods, thus providing additional information about species distribution dynamics that can be further used to set valid demographical hypotheses (Richards et al., 2007; Collevatti et al., 2013; **Figure 1A**). ENMs can actually generate multiple and independent hypotheses of species distribution history that reflect, at same time, ecological, and biogeographical realism (Collevatti et al., 2013). This is because different assumptions about the dynamics of species' ecological niches can be explored using different ENM methods (algorithms) and palaeoclimatic simulations, which are in turn based on different modeling assumptions or types of training data (modeling uncertainty). Fossil records, when available, can be used to improve predictions of past distributions (either by providing

additional information about species environmental preferences or by validating ENM predictions; Lorenzen et al., 2011; Varela et al., 2011), or to propose additional demographic hypotheses not discriminated from ENMs (**Figure 1A**).

Moreover, ENMs also provide the environmental context of population movements across geographical space through time (**Figure 1B**), as proposed by Collevatti et al. (2012b). Dispersal events may be inferred from spatially explicit analyses of population genetic structure in relation to fluctuations of climatic suitability through time (a direct measure from ENMs in areas where populations were sampled), space (habitat tracking or range shifts observed from consensual palaeodistribution as predicted by ENMs), and location of historical refugia (areas climatically suitable for the focal species throughout the time).

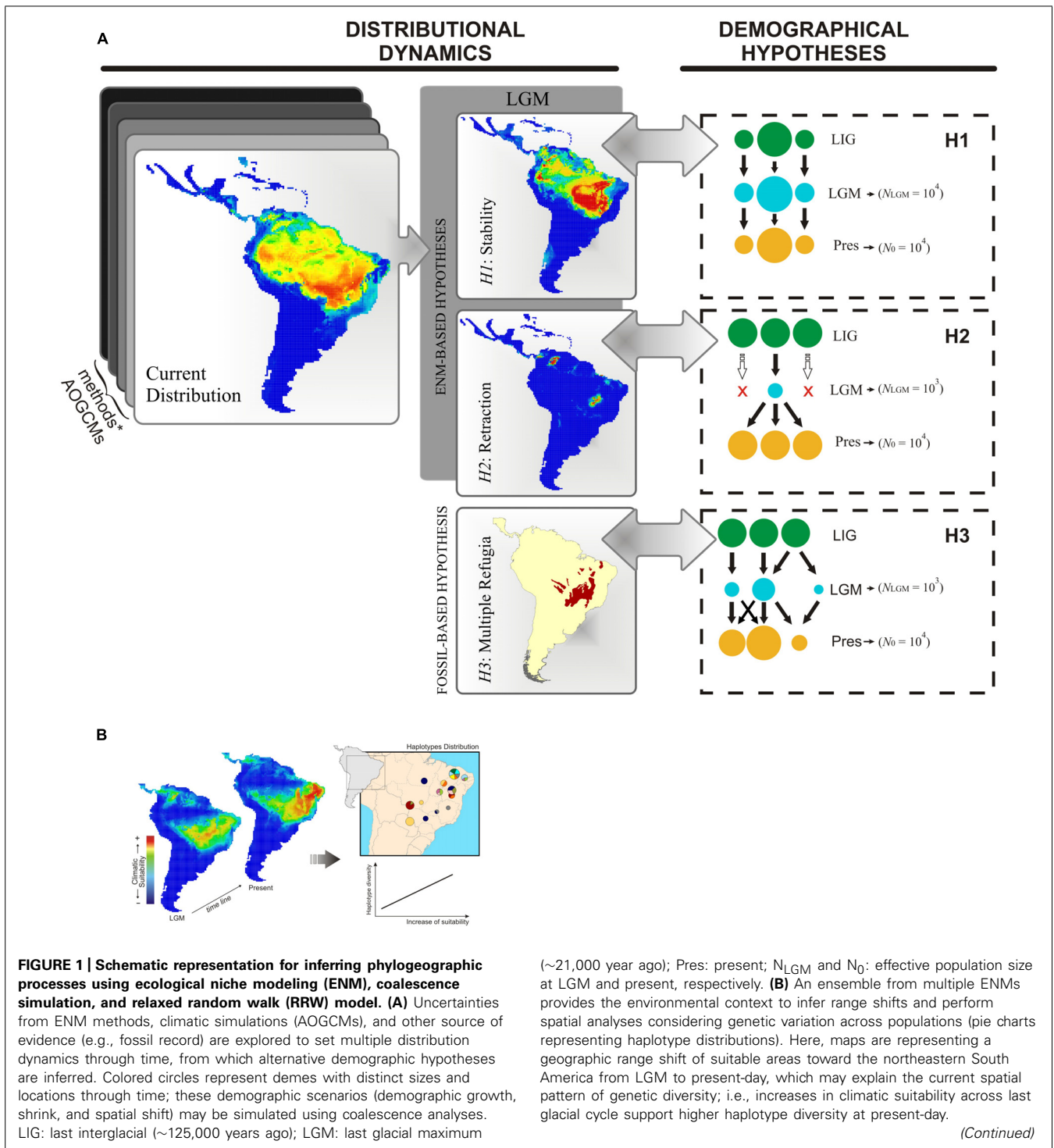
However, today's practices of coupling ENMs with phylogeographic data analysis are not free of caveats. Currently, most climatic reconstructions through Atmosphere-Oceanic General Circulation Models (AOGCMs) are mainly available for short time slices, as in the case of the last 21,000 years, i.e., the time interval since the last glacial maximum (see the most recent palaeoclimatic simulations in PMIP3¹ – and CMIP5 databases²). Additionally, the major lineage divergences and common phylogeographic breaks among species may have occurred earlier than this. In fact, coalescent analyses infer the time to the most recent common ancestor (TMRCA) in a time scale derived from the molecular sequence substitution rate under a particular demographic model (Kuhner, 2008). As a consequence, molecular sequences with lower mutation rates (such as chloroplast DNA for example) may lead to older divergence dating compared to the time interval of ENMs. Consequently, the predictions from ENMs and coalescent analyses would be temporarily discordant. We propose two solutions to this apparent weakness.

First, if palaeodistribution modeling is based only on climatic conditions and, most importantly, if predictions are intended to test the genetic legacy from recurrent glacial cycles, the modeler may set the distribution dynamics across the last glacial cycle (for which palaeoclimatic simulations are commonly available) and assumes analogous dynamics through the older cycles. Although separate glacial cycles have provided idiosyncratic dynamics on small temporal and geographical scales, the general pattern of intermittent glacial and interglacial periods was common throughout the Quaternary. Thus, considering broad geographical scales, it seems acceptable to assume that similar distributional dynamics have occurred across different Quaternary glaciations. In contrast, ENMs could be projected for deeper periods (e.g., last interglacial – 125,000 year – mid-Pliocene – ~3 Million years), avoiding such important assumption, although few AOGCMs are currently available for such periods (see Stone et al., 2013, and the special issue “PlioMIP: experimental design, mid-Pliocene boundary conditions and implementation” at the journal *Geoscientific Model Development*, available at³). Thus, it is important that the assumption of analogous dynamics through time is validated by comparing the potential population movements across

¹<https://pmip3.lsce.ipsl.fr/>

²<http://cmip-pcmdi.llnl.gov/>

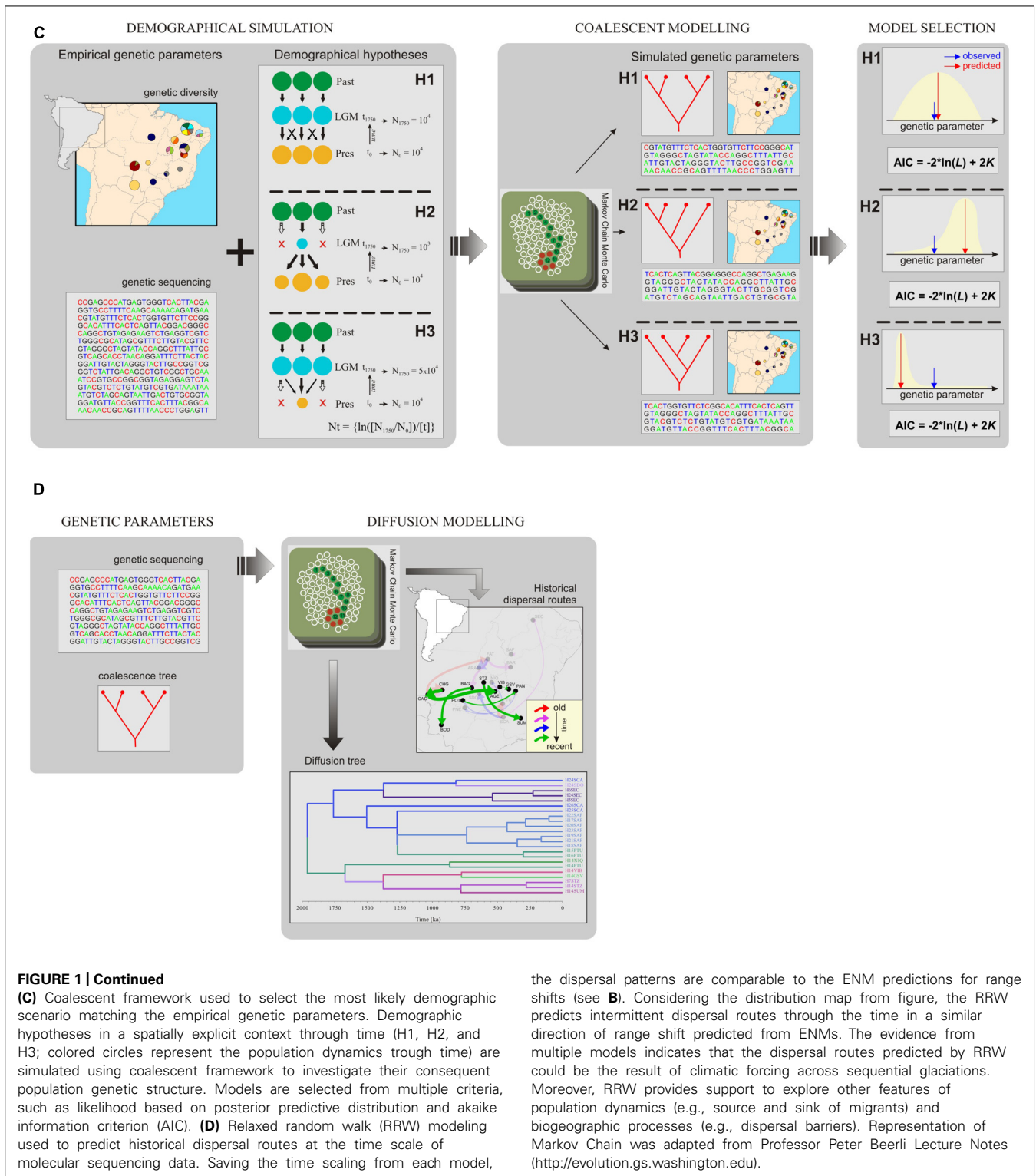
³http://www.geosci-model-dev.net/special_issue5.html



geographical space including the specific time slice from ENM predictions with patterns of lineage diffusion explicitly simulated by relaxed random walk (RRW) model, which encompasses deeper time that is proportional to the molecular evolution of the sequence used (Lemey et al., 2009, 2010; see also the section below). If the general patterns of population dispersal during cooling and warming phases are concordant between ENM and RRW,

then assuming similar distributional dynamics across Quaternary glaciations (on broad spatio-temporal scales) is not equivocal (but see next proposal).

Second, due to recent advances on Pliocene–Pleistocene stacked estimates of isotopic globally distributed oxygen (e.g., Lisiecki and Raymo, 2005), climatic conditions may be extended backward by using the climate change between LGM and present-day to



interpolate climate trends to older glacial cycles following the proportional oscillation across the deeper oxygen curve. At the same time, if predictions are at species level and comparative across multiple species, the ancestral state of the species niche may still be simulated across a phylogenetic hypothesis, so that

palaeodistributions are automatically obtained at the deeper time (see example in Lawing and Polly, 2011; Rödder et al., 2013). When possible, these solutions should be preferred instead of assuming similar distribution dynamics across different Quaternary glaciations. Palaeoclimatic simulations for deeper times may improve

our approach by projecting ENM predictions directly or using the temporal interpolation.

DEMOGRAPHIC HISTORY SIMULATION AND MODEL SELECTION

To trace demographic history, demographic scenarios can be modeled, and simulated under a coalescent framework (Figure 1C; Kingman, 1982; see Csilléry et al., 2010 for a review of available software). Briefly, the available software runs independent simulations for each sequence region based on demographic parameters such as migration and effective population size, and under a given evolutionary model, sequence length, and mutation rate. Usually, simulation output includes genetic diversity estimates such as haplotype and nucleotide diversities or expected heterozygosity under Hardy–Weinberg equilibrium for genotypic data, number of haplotypes or alleles, parameters for neutrality, and demographic expansion tests, sequences, and genotypes.

The alternative models may be compared using several criteria (Figure 1C). For instance, the posterior estimates of genetic parameters for the alternative demographic scenarios can be compared with the empirical haplotype and nucleotide diversity (Csilléry et al., 2010). The likelihood of each model can be obtained from the posterior predictive distribution and the alternative models can be compared using the Akaike Information Criterion (AIC; see Burnham and Anderson, 2002). Model fitting may also be performed generating coalescent trees under each simulated demographic scenario and compared with observed coalescence time using ABC implemented in MTML-msBayes (Huang et al., 2011).

However, despite multiple lines of evidence to design alternative demographic hypotheses, spatially explicit modeling is yet to be developed. Although some advances have been made with software like SPLATCHE2 (Ray et al., 2010) and PHYLOGEOSIM 1.0⁴ (Dellicour et al., 2014), more complex palaeodistribution dynamics that can differentiate among some predictions are still unavailable. For instance, modeling demographic scenarios for range shift and range expansion is still a challenge because both scenarios may result from similar demographical dynamics (smaller effective population size in the past than in the present-day), but generate different genetic signatures due to spatial context (see Excoffier et al., 2009). Also, different range shift scenarios may generate distinct genetic signatures depending on the spatial direction of colonization of founding lineages (Excoffier et al., 2009; Waters et al., 2013).

DIFFUSION MODEL AND CLUES TO DISPERSAL ROUTES

The lack of fossil records for most species makes understanding historical dispersal routes difficult especially in Neotropics. Thus, integrating direct spatio-temporal reconstruction of lineage dispersal may give new insights on the pathway of lineage dispersal to better understand phylogeographic patterns.

Lemey et al. (2009, 2010) proposed a Bayesian statistical approach to infer continuous phylogeographic diffusion using a RRW model, while simultaneously reconstructing the evolutionary history in time from molecular sequence data (Figure 1D).

More specifically Lemey et al.'s (2009) approach describes the phylogeographical diffusion processes by stochastically selecting a diffusion rate scalar on each branch of the rooted phylogeny from an underlying discretized rate distribution while running a Bayesian Markov Chain Monte Carlo model. Consequently, two important advantages arise from Lemey et al.'s (2009) approach: (1) relaxing the most restrictive assumption of the standard Brownian diffusion model, and (2) inferring the migration process in natural time scales (i.e., the time scale of the molecular sequence substitution process). Moreover, because this framework is based on stochastic models, it naturally accesses the uncertainties along the ancestral state reconstructions and the underlying phylogeographic process (Lemey et al., 2009, 2010), an essential component from any multi-model inference approach (Millington and Perry, 2011). We understand that this approach may be a fine complement to the static ENM predictions of population dispersal, using explicitly simulated dispersal routes in the evolutionary time scale of molecular sequences. Thus, it will become an indispensable component in a multi-model framework for phylogeographical inferences. RRW also deals with uncertainties along the ancestral state reconstructions and the underlying phylogeographical process because it is based on stochastic models (Lemey et al., 2009, 2010). The relaxed random walk model is implemented in the software BEAST 1.8.0 (Drummond and Rambaut, 2007) that analyses sequence evolution, demographic model, and lineage diffusion in space and time simultaneously, and the spatio-temporal reconstruction can be performed using SPREAD 1.0.6 (Bielejec et al., 2011). Although a promising approach, we understand that it is still necessary to find explicit methods to couple predictions from ENM and RRW.

EXPANDING THE MULTI-MODEL FRAMEWORK FOR COMPARATIVE PHYLOGEOGRAPHY

Following this reasoning of model-based inference, we propose the extension of the framework coupling ENM, coalescent simulation and the RRW model for comparative phylogeography (Figure 2). In a nutshell, for such an upgrade, alternative demographical hypotheses should firstly be set considering the complete range of distribution dynamics considering all analyzed species. In a similar manner to statistical phylogeography, all uncertainties from ENMs, and other sources of palaeodistribution scenarios should be explored, such as palaeovegetation reconstruction and fossil records, to set individual species dynamics.

However, species may share common palaeodistribution dynamics or not, under ENM predictions. Due to the differences in life-history or functional traits, species may have responded differently to climate changes or other biogeographical processes (Colinvaux et al., 2000). For instance, *Tabebuia impetiginosa* (Collevatti et al., 2012a) and *Astronium urundeuva* (Caetano et al., 2008), both from seasonally dry forests in Brazil, expanded their range in response to drier and cooler periods of the glacial cycles in Neotropics, whereas in Brazilian savannas, *Caryocar brasiliense* (Collevatti et al., 2012b), and *T. aurea* (Collevatti et al., 2014) showed population retraction in multiple refugia as response to the same climatic events. The lack of common palaeodistribution dynamics, however, does not mean that different historical biogeographical process affected each species

⁴<http://ebe.ulb.ac.be/ebe/Software.html>

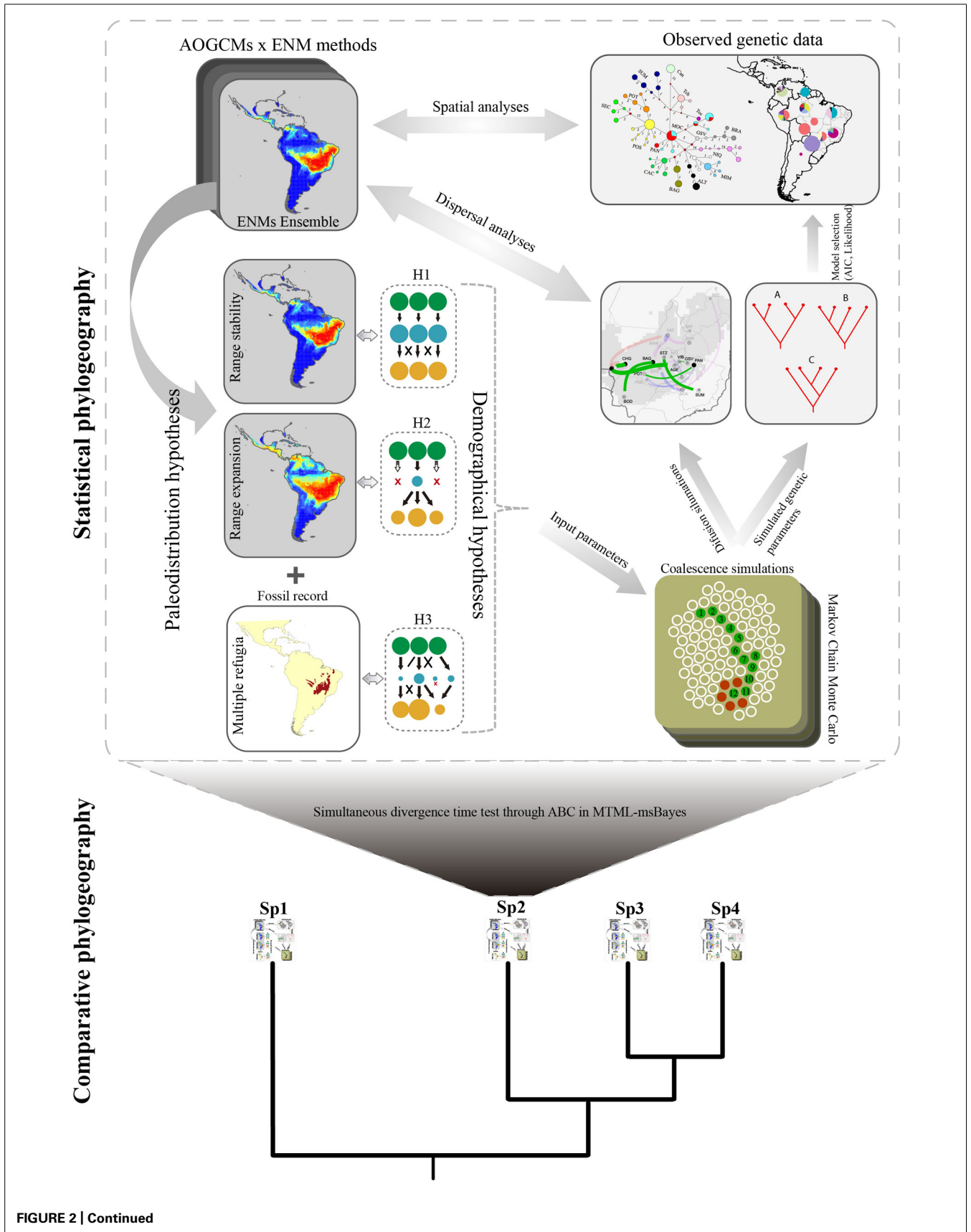


FIGURE 2 | Continued

FIGURE 2 | Continued**Schematic view of the generalized framework for coupling ecological niche modeling, coalescent simulation, and diffusion model in a statistical comparative phylogeography approach.**

Palaeodistribution maps resulting from ENMs are used to generate demographic hypotheses (see **Figure 1**) in a spatially explicit context through time (hypotheses H1 and H2; colored circles represent the population dynamics), according to the range shifts in species distributions in the past (e.g., range retraction or range stability). Other hypotheses may also be set based on fossil records (hypothesis H3) or *a priori* biogeographic scenario. Uncertainties in setting alternative hypotheses can be incorporated into the framework using several ENM methods and projecting distributions through different AOGCMs. In a next step, simulated coalescence structures are compared with observed data in a model selection approach, allowing selecting among the demographic hypothesis (H1, H2, or H3) the most likely to generate the current phylogeographic structure derived from molecular data. At the same time, phylogeographic diffusion models allow reconstructing colonization routes that are compared with palaeodistribution maps. Finally, this framework can be expanded into a multi-species comparative approach allowing inferring how whole assemblages responded to the interplay between climate changes, geographic barriers, and demographic processes, shaping the current patterns of species distribution, and biodiversity. Coalescent time for each species can be compared using, for instance, Approximate Bayesian Computation implemented in MTML-msBayes. Representation of Markov Chain was adapted from Professor Peter Beerli Lecture Notes (<http://evolution.gs.washington.edu>).

if evolutionary timing matches. Divergence timing and demographic response may be compared to better understand how Quaternary glaciations affected multiple species from distinct regions and with different traits. Concerning the comparative analyses, demographic hypotheses should be set *a priori*, and therefore biogeographical hypotheses would be particularly investigated for species with unique characters; i.e., hypotheses may be usually proposed for entire biomes or biotas (e.g., Pleistocene Arc hypothesis for Neotropical seasonally dry forests, see Prado and Gibbs, 1993) and thus may be simulated for all species from the same biome in a comparative phylogeography framework (**Figure 2**).

Consequently, the coalescent simulations based on the palaeodistribution scenarios may also be performed for all species and compared among species from the same functional group, similar ecosystems, or with similar life-histories (e.g., similar pollination and dispersal syndromes). The role of life-history or quantitative traits in shaping general phylogeographic patterns may ultimately be investigated using random or mixed effects models in meta-regression, weighting evidence by its level of uncertainty (Stanley and Jarrell, 1989). Whatever the source, the higher the uncertainty for a species (e.g., from ENM predictions) the lower is its influence to draw general phylogeographical inference under this comparative framework.

Moreover, the reconstruction of colonization routes would complement the understanding of how unique historical process affected multiple species in a broad biogeographical hypothesis (e.g., Taberlet et al., 1998). Even with a higher number of species studied, understanding the role of vicariance and dispersal routes is compromised in most regions worldwide because of the lack of direct empirical evidence from fossil records at community level. Thus, integrating direct spatio-temporal reconstruction of lineage diffusion with ecological niche modeling and coalescent simulation may indicate the pathways where multiple lineages have dispersed and their genetic legacies as a response to Quaternary

climate changes and other biogeographic processes. For *T. aurea*, for instance, reconstruction of colonization routes unraveled the role of populations with higher genetic diversity at the edge of the historical climatic refugium as a source of migrants, whereas populations at the center of climatically stable areas worked usually as a sink of migrants (Collevatti et al., submitted).

In addition, integrating direct spatio-temporal reconstruction of lineage diffusion with dispersal routes predicted by the fossil record may allow validation and improvement of the lineage diffusion model. For instance, Lima et al. (2014) used the pollen fossil record of *Mauritia flexuosa*, a Neotropical swamp palm, to validate the predictions of ENM on population range shifts. The comparison of dispersal routes based on RRW models with pollen fossil records and ENM predictions can be applied to predict and validate dispersal routes during spatial population displacements.

In conclusion, along with the flexible and integrative nature of our multi-model framework in the context of the statistical phylogeography, its expansion in a comparative direction also makes it comprehensive. This aspect of our multi-model inference framework is particularly useful to investigate the complex dynamics and current patterns of genetic diversity in response to processes operating on multiple taxonomic levels as approached in comparative phylogeography.

ACKNOWLEDGMENTS

Our research program integrating macroecology and molecular ecology has been continuously supported by grants to the research network GENPAC (Geographical Genetics and Regional Planning for natural resources in Brazilian Cerrado) supported by CNPq/MCT/CAPES/FAPEG (projects no. 564717/2010-0, 563727/2010-1 and 563624/2010-8), and the network Rede Cerrado CNPq/PPBio (project no. 457406/2012-7) that we gratefully acknowledge.

REFERENCES

- Avice, J. C. (2000). *Phylogeography: The History and Formation of Species*. Cambridge: Harvard University Press.
- Bell, R. C., MacKenzie, J. B., Hickerson, M. J., Chavarria, K. L., Cunningham, M., Williams, S., et al. (2012). Comparative multi-locus phylogeography confirms multiple vicariance events in co-distributed rainforest frogs. *Proc. Biol. Sci.* 279, 991–999. doi: 10.1098/rspb.2011.1229
- Bermingham, E., and Moritz, C. (1998). Comparative phylogeography: concepts and applications. *Mol. Ecol.* 7, 367–369. doi: 10.1046/j.1365-294x.1998.00424.x
- Bielejec, E., Rambaut, A., Suchard, M. A., and Lemey, P. (2011). SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* 27, 2910–2912. doi: 10.1093/bioinformatics/btr481
- Burnham, K. P., and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach*. New York: Springer.
- Caetano, S., Prado, D., Pennington, R. T., Beck, S., Oliveira-Filho, A., Spichiger, R., et al. (2008). The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Mol. Ecol.* 17, 3147–3159. doi: 10.1111/j.1365-294X.2008.03817.x
- Carnaval, A. C., and Moritz, C. (2008). Historical climate modeling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *J. Biogeogr.* 35, 1187–1201. doi: 10.1111/j.1365-2699.2007.01870.x
- Carstens, B. C., and Knowles, L. L. (2010). Navigating the unknown: model selection in phylogeography. Models of population structure: tools for thinkers. *Mol. Ecol.* 19, 4581–4582. doi: 10.1111/j.1365-294X.2010.04851.x
- Chan, L. M., Brown, J. L., and Yoder, A. D. (2011). Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Mol. Phylog. Evol.* 59, 523–537. doi: 10.1016/j.ympev.2011.01.020

- Collevatti, R. G., Lima-Ribeiro, M. S., Terribile, L. C., Rosa, F. F., Guedes, L. B. S., and Telles, M. P. C. (2014). Recovering species demographic history from multi-model inference: the case of a Neotropical savanna tree species. *BMC Evol. Biol.* 14:213. doi: 10.1186/s12862-014-0213-0
- Collevatti, R. G., Terribile, L. V., Lima-Ribeiro, M. S., Nabout, J. C., Oliveira, G., Rangel, T. F., et al. (2012a). A coupled phylogeographic and species distribution modeling approach recovers the demographic history of a Neotropical seasonally dry forest tree species. *Mol. Ecol.* 21, 5845–5863. doi: 10.1111/mec.12071
- Collevatti, R. G., Lima-Ribeiro, M. S., Souza-Neto, A. C., Franco, A. A., and Terribile, L. V. (2012b). Recovering the demographical history of a Brazilian Cerrado tree species *Caryocar brasiliense*: coupling ecological niche modeling and coalescent analyses. *Nat. Conserv.* 10, 169–176. doi: 10.4322/natcon.2012.024
- Collevatti, R. G., Terribile, L. V., Oliveira, G., Lima-Ribeiro, M. S., Nabout, J. C., Rangel, T. F., et al. (2013). Drawbacks to palaeodistribution modelling: the case of South American seasonally dry forests. *J. Biogeogr.* 40, 345–358. doi: 10.1111/jbi.12005
- Colinvaux, P. A., Oliveira, P. E., and Bush, M. B. (2000). Amazonian and Neotropical plant communities on glacial time-scales: the failure of the aridity and refuge hypotheses. *Quat. Sci. Rev.* 19, 141–169. doi: 10.1016/S0277-3791(99)00059-1
- Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., and François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends Ecol. Evol.* 25, 410–418. doi: 10.1016/j.tree.2010.04.001
- Dellicour, S., Kastally, C., Hardy, O. J., and Mardulyn, P. (2014). Comparing phylogeographic hypotheses by simulating DNA sequences under a spatially explicit model of coalescence. *Mol. Biol. Evol.* 31, 3359–3372. doi: 10.1093/molbev/msu277
- Drummond, A. J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/1471-2148-7-214
- Excoffier, L., Foll, M., and Petit, R. J. (2009). Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* 40, 481–501. doi: 10.1146/annurev.ecolsys.39.110707.173414
- Grueber, C. E., Nakagawa, S., Laws, R. J., and Jamieson, I. G. (2011). Multi-model inference in ecology and evolution: challenges and solutions. *J. Evol. Biol.* 24, 699–711. doi: 10.1111/j.1420-9101.2010.02210.x
- Huang, W., Takebayashi, N., Qi, Y., and Hickerson, M. J. (2011). MTML-msBayes: approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC Bioinform.* 12:1. doi: 10.1186/1471-2105-12-1
- Kingman, J. F. C. (1982). The coalescent. *Stoch. Proc. Appl.* 13, 235–248. doi: 10.1016/0304-4149(82)90011-4
- Knowles, L. L. (2009). Statistical phylogeography. *Annu. Rev. Ecol. Evol. Syst.* 40, 593–612. doi: 10.1146/annurev.ecolsys.38.091206.095702
- Knowles, L. L., Carstens, B. C., and Keat, M. L. (2007). Coupled genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Curr. Biol.* 17, 1–7. doi: 10.1016/j.cub.2007.04.033
- Knowles, L. L., and Maddison, W. P. (2002). Statistical phylogeography. *Mol. Ecol.* 11, 2623–2635. doi: 10.1046/j.1365-294X.2002.01637.x
- Kuhner, M. K. (2008). Coalescent genealogy samplers: windows into population history. *Trends Ecol. Evol.* 24, 86–93. doi: 10.1016/j.tree.2008.09.007
- Lagerholm, V. K., Sandoval-Castellanos, E., Ehrlich, D., Abrahamson, N. I., and Nadachowski, A. (2014). On the origin of the Norwegian lemming. *Mol. Ecol.* 23, 2060–2071. doi: 10.1111/mec.12698
- Lawing, A. M., and Polly, P. D. (2011). Pleistocene climate, phylogeny, and climate envelope models: an integrative approach to better understand species' response to climate change. *PLoS ONE* 6:e28554. doi: 10.1371/journal.pone.0028554
- Lemey, P., Rambaut, A., Drummond, A. J., and Suchard, M. A. (2009). Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* 5:e1000520. doi: 10.1371/journal.pcbi.1000520
- Lemey, P., Rambaut, A., Welch, J. J., and Suchard, M. A. (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* 27, 1877–1885. doi: 10.1093/molbev/msq067
- Lima, N. E., Lima-Ribeiro, M. S., Tinoco, C. F., Terribile, L. C., and Collevatti, R. G. (2014). Phylogeography and ecological niche modelling, coupled with the fossil pollen record, unravel the demographic history of a Neotropical swamp palm through the Quaternary. *J. Biogeogr.* 41, 678–686. doi: 10.1111/jbi.12269
- Lisiecki, L. E., and Raymo, M. E. (2005). A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleoceanography* 20, PA1003. doi: 10.1029/2004PA001071
- Lorenzen, E. D., Nogués-Bravo, D., Orlando, L., Weinstock, J., Binladen, J., Marske, K. A., et al. (2011). Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* 479, 359–364. doi: 10.1038/nature10574
- Martínez-Meyer, E., Peterson, A. T., and Hargrove, W. W. (2004). Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Global Ecol. Biogeogr.* 13, 305–314. doi: 10.1111/j.1466-822X.2004.00107.x
- Metcalf, J. L., Prost, S., Nogués-Bravo, D., DeChaine, E. G., Anderson, C., Batra, P., et al. (2014). Integrating multiple lines of evidence into historical biogeography hypothesis testing: a *Bison bison* case study. *Proc. R. Soc. B* 281, 2013–2782. doi: 10.1098/rspb.2013.2782
- Millington, J. D. A., and Perry, G. L. W. (2011). Multi-model inference in biogeography. *Geo. Compass* 5/7, 448–463. doi: 10.1111/j.1749-8198.2011.00433.x
- Nogués-Bravo, D. (2009). Predicting the past distribution of species climatic niches. *Global Ecol. Biogeogr.* 18, 521–531. doi: 10.1111/j.1466-8238.2009.00476.x
- Paul, C. R. C. (2009). The fidelity of the fossil record: the improbability of preservation. *Palaeontology* 52, 485–489. doi: 10.1111/j.1475-4983.2009.00872.x
- Prado, D. E., and Gibbs, P. E. (1993). Patterns of species distributions in the dry seasonal forests of South America. *Ann. Missouri Bot. Gard.* 80, 902–927. doi: 10.2307/2399937
- Ray, N., Currat, M., Foll, M., and Excoffier, L. (2010). SPLATCHE2: a spatially-explicit simulation framework for complex demography, genetic admixture and recombination. *Bioinformatics* 26, 2993–2994. doi: 10.1093/bioinformatics/btq579
- Richards, C. L., Carstens, B. C., and Knowles, L. L. (2007). Distribution modeling and statistical phylogeography: an integrative framework for testing biogeographic hypotheses. *J. Biogeogr.* 34, 1833–1845. doi: 10.1111/j.1365-2699.2007.01814.x
- Rödger, D., Lawing, A. M., Flecks, M., Ahmadzadeh, F., Dambach, J., Engler, J. O., et al. (2013). Evaluating the significance of paleophylogeographic species distribution models in reconstructing Quaternary range-shifts of Nearctic chelonians. *PLoS ONE* 8:e72855. doi: 10.1371/journal.pone.0072855
- Stanley, T. D., and Jarrell, S. B. (1989). Meta-regression analysis: a quantitative method of literature surveys. *J. Econ. Surv.* 3, 161–170. doi: 10.1111/j.1467-6419.1989.tb00064.x
- Stephens, P. A., Buskirk, S. W., and Rio, C. M. (2006). Inference in ecology and evolution. *Trends Ecol. Evol.* 22, 192–197. doi: 10.1016/j.tree.2006.12.003
- Stone, E. J., Bakker, P., Charbit, S., Ritz, S. P., and Varma, V. (2013). A climate model inter-comparison of last interglacial peak warmth. *Past Glob. Chang.* 21, 32–33.
- Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., and Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* 7, 453–464. doi: 10.1046/j.1365-294x.1998.00289.x
- Varela, S., Lobo, J. M., and Hortal, J. (2011). Using species distribution models in paleobiogeography: a matter of data, predictors and concepts. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 310, 451–463. doi: 10.1016/j.palaeo.2011.07.021
- Velleman, P. F. (2008). Truth, damn truth, and statistics. *J. Stat. Educ.* 16, 1–14.
- Waters, J. M., Fraser, C. I., and Hewitt, G. M. (2013). Founder takes all: density-dependent processes structure biodiversity. *Trends Ecol. Evol.* 28, 78–85. doi: 10.1016/j.tree.2012.08.024

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 June 2014; accepted: 22 January 2015; published online: 17 February 2015.

Citation: Collevatti RG, Terribile LC, Diniz-Filho JAF and Lima-Ribeiro MS (2015) Multi-model inference in comparative phylogeography: an integrative approach based on multiple lines of evidence. *Front. Genet.* 6:31. doi: 10.3389/fgene.2015.00031

This article was submitted to *Evolutionary and Population Genetics*, a section of the journal *Frontiers in Genetics*.

Copyright © 2015 Collevatti, Terribile, Diniz-Filho and Lima-Ribeiro. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.